Husimi coordinates of multipartite separable states

Georges Parfionov and Romàn R. Zapatrin

Informatics Dept., The State Russian Museum,

Inżenernaya 4, 191186, St.Petersburg, Russia;

e-mail: Roman.Zapatrin at Gmail.com

April 27, 2010

Abstract

A parametrization of multipartite separable states in a finite-dimensional Hilbert space is suggested. It is proved to be a diffeomorphism between the set of zero-trace operators and the interior of the set of separable density operators. The result is applicable to any tensor product decomposition of the state space. An analytical criterion for separability of density operators is established in terms of the boundedness of a sequence of operators.

Introduction

We explore multipartite quantum systems with state space $H^{(1)} \otimes \cdots \otimes H^{(N)}$, assuming the factors are Hilbert spaces of *finite* dimension:

$$\dim H^{(1)}, \ldots, \dim H^{(N)} < \infty$$

A pure state of the system is a state whose density operator is one-dimensional projector. Among pure states we there states of special kind:

$$\mathrm{P} = P^{(1)} \otimes \cdots \otimes P^{(N)}$$

where $P^{(k)}$ is a one-dimensional projector in $H^{(k)}$. A density operator ρ representable as a convex combination

$$\rho = \sum w_i \mathrm{P}_i, \qquad \mathrm{P}_i \in \mathbf{P}$$

where $\mathbf{P}$ is the set of all pure product states, is said to be separable. Such description of the set of separable density operators is not constructive as their appropriate representation is *not unique*. In this paper a uniform representation of separable density operators ρ in the form

$$\rho \;=\; \frac{1}{Z}\int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(\hat{\beta}\cdot \mathrm{P})}\mathrm{P}\mathrm{d}\sigma(\mathrm{P})$$

is suggested. This representations turns out to be *unique*, therefore, it can be treated as coordinatization of the set of separable density operators. Along these lines, a criterion to determine the separability of a given density operator ρ is formulated.

The paper is organized as follows. In Section 1 separable density operators are treated as barycenters of *continuously* distributed unit masses on the set $\mathbf{P}$ of pure product states. This gives rise to the equation (5) on which the coordinatization of separable density operators is based. In Section 2 a scalar function (6) is studied discriminating the separability in many-particle setting. This makes it possible to formulate a criterion of multipartite separability in terms of boundedness of certain sequence of operators. In Section 3 the range of the suggested coordinatization of separable density operators is determined. The range turns out to be the interior of the set of separable density operators, and the coordinatization is proved to be a diffeomorphism.

1 Separable density operators viewed as barycenters

In this section we describe how density operators in a finite-dimensional Hilbert space can be described as barycenters of continuous probability distributions on a set of one-dimensional projectors.

1.1 From finite sums to continuous distributions

We study a multipartite quantum system whose state space $\mathbf{H}$ is a finite tensor product of finite-dimensional Hilbert spaces $\mathbf{H} = H^{(1)} \otimes \cdots \otimes H^{(N)}$. A

density operator ρ in $\mathbf{H}$ is called *separable* whenever it is an element of the convex hull of the set $\mathbf{P}$ of pure product states. Since $\mathbf{P}$ is the set of extreme points of the set of all separable density operators, Choquet theorem makes it possible to write down any such ρ as an integral

$$\rho = \int_{\mathbf{P}} \mathrm{P} \mathrm{d}\mu(\mathrm{P}) \tag{1}$$

over certain probability measure μ on $\mathbf{P}$. In other words, we may represent ρ as the *barycenter* of certain distribution of unit mass on the set $\mathbf{P}$. Such representation is known to be essentially *non-unique*. According to Carathéodory theorem this mass may have discrete finite distribution

$$\rho = \sum w_\alpha \mathrm{P}_\alpha$$

with the measure $\mu = \sum w_\alpha \delta_{\mathrm{P}_\alpha}$ being a sum of atomic measures $\delta_{\mathrm{P}_\alpha}$.

We emphasize that representation (1) also comprises *continuous* distributions

$$\rho \;=\; \int_{\mathbf{P}} \mathrm{P} w(\mathrm{P}) \mathrm{d}\sigma(\mathrm{P}) \tag{2}$$

where $w(\mathrm{P})$ is a positive continuous function and the integration is carried out over the probability measure $\mathrm{d}\sigma$ *invariant* with respect to all *local* unitary transformations in $\mathbf{H}$.

So, a density operator ρ in the product space $\mathbf{H} = H^{(1)} \otimes \cdots \otimes H^{(N)}$ is separable when it is a barycenter of a continuous distribution $w(\mathrm{P})$ on the set $\mathbf{P}$ of pure product vectors. In other words, that means that a function $w(\mathrm{P})$ exists satisfying the equation

$$\int_{\mathbf{P}} \mathrm{P} w(\mathrm{P}) \mathrm{d}\sigma(\mathrm{P}) \;-\; \rho \;=\; 0 \tag{3}$$

1.2 Main equation

In this section we confine ourselves to distributions $w(\mathrm{P})$ of specific form and explore the existence of appropriate solutions.

Let us first try to find solutions of (3) of the form

$$w(\mathrm{P}) = \mathrm{e}^{\mathrm{Tr}(B\mathrm{P})} \tag{4}$$

where B is an Hermitian operator. Then solving the equation (3) reduces to finding an operator X satisfying what we call *main equation*

$$\int_{\mathbf{P}} \mathrm{P}\mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \mathrm{d}\sigma(\mathrm{P}) \;-\; \rho \;=\; 0 \tag{5}$$

In order to verify the existence of a solution of (5), introduce the function

$$G(X) \;=\; \int_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \mathrm{d}\sigma(\mathrm{P}) \;-\; \mathrm{Tr}(X\rho) \tag{6}$$

whose gradient is

$$\nabla G(X) \;=\; \int_{\mathbf{P}} \mathrm{P}\mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \mathrm{d}\sigma(\mathrm{P}) \;-\; \rho \tag{7}$$

So, solving (5) reduces to finding an extremum of the function $G(X)$. Taking into account that $G(X)$ is *convex*, the solution of (5) exists only when the *minimum* of $G(X)$ exists. In particular, when ρ is entangled, there is no way for the function $G(X)$ to have minimum:

Proposition 1. *If the density operator ρ is entangled, then*

$$\inf G(X) \;=\; -\infty$$

Proof. The set of all separable states is closed, therefore, if ρ is not separable, there exists a hyperplane, defined by a self-adjoint operator X such that $\forall \mathrm{P} \in \mathbf{P} \quad \mathrm{Tr}(\mathrm{P}X) < 0$, while $\mathrm{Tr}(\rho X) > 0$. Denote

$$\begin{aligned} a &= \mathrm{Tr}(\rho X) \\ b &= \max \mathrm{Tr}(\mathrm{P}\, X) \end{aligned}$$

Then $a > 0, b < 0$, so $\mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \leq \mathrm{e}^{kb}$ and

$$G(kX) \;\leq\; \int \mathrm{e}^{kb} \mathrm{d}\sigma(\mathrm{P}) - ka \;\to\; -\infty$$

as $k \to \infty$. □

So far, finding conditions for the existence of minimum becomes essential to judge if ρ is entangled or not.

2 Exponential distributions and their approximations

2.1 An interlude on minima of convex functions

Proposition 2. *Let a convex function $F(X)$ has minimum on a finite-dimensional space $\mathcal{E}$ and the minimal point is unique. Then*

$$\lim_{X\to\infty} F(X) = +\infty \tag{8}$$

Proof. With no loss of generality assume the minimum to be attained at 0, then $F(X) > F(0)$ for any $X \neq 0$. Suppose $\lim_{X\to\infty} F(X) \neq +\infty$, then there exists such M and such sequence $X_k \to \infty$ that $F(X_k) \leq M$. Then

$$\lambda_k = \frac{1}{\|X_k\|} \to 0$$

Denoting

$$E_k \;=\; \frac{X_k}{\|X_k\|} \;=\; \lambda_k X_k$$

we obtain a bounded sequence in the finite-dimensional space $\mathcal{E}$, which contains a converging subsequence. With no loss of generality denote this subsequence E_k and its limit

$$E = \lim E_k, \quad \|E\| = 1$$

Then

$$E_k \;=\; \lambda_k X_k \;=\; (1-\lambda_k)\,0 + \lambda_k X_k$$

so, due to the convexity of F

$$F(E_k) \leq (1-\lambda_k)\,F(0) + \lambda_k F(X_k)$$

therefore

$$F(E_k) \leq (1-\lambda_k)\,F(0) + \lambda_k M$$

for sufficiently large k. Taking the limit and recalling that $\lambda_k \to 0$, and using the continuity of convex functions (see, e.g. [1]), we come to the contradiction: $F(E) \leq F(0)$. $\square$

So far, we obtained a necessary condition for the unique minimum of a convex function to exist. This condition (8) is also sufficient for a minimum (not necessarily strict) to exit. In the sequel we shall need a more verifiable sufficient condition:

Proposition 3. *Let $F(X)$ be a convex function on a finite-dimensional space $\mathcal{E}$. If*

$$\forall\, X \neq 0 \qquad \lim_{t\to\infty} F(tX) = +\infty \tag{9}$$

then there exists minimum of $F(X)$ on $\mathcal{E}$.

Proof. First prove that

$$\lim_{X\to\infty} F(X) = +\infty$$

Suppose this is not the case, then there exists a number M and a sequence $X_k \to \infty$ such that $F(X_k) \leq M$ for all k. Denoting

$$E_k \;=\; \frac{X_k}{\|X_k\|}$$

we obtain a bounded sequence in the finite-dimensional space $\mathcal{E}$, which contains a converging subsequence. With no loss of generality denote this subsequence E_k and its limit

$$E = \lim E_k, \quad \|E\| = 1$$

Now take an arbitrary $t \geq 0$ and N such that $\|X_k\| \geq t$ for all $k \geq N$. Then $\|tE_k\| = t \leq \|X_k\|$. Denote

$$\lambda_k = \frac{t}{\|X_k\|}$$

then $0 \leq \lambda_k \leq 1$ and $tE_k = \lambda_k X_k$. We may treat tE_k as convex combination

$$tE_k = (1-\lambda_k)\,0 + \lambda_k X_k$$

and apply Jensen's inequality

$$F(tE_k) \leq (1-\lambda_k)F(0) + \lambda_k F(X_k) \leq F(0) + M$$

for all $k \geq N$. Taking the limit, we get $F(tE) \leq F(0) + M$, which contradicts with the statement $F(tE) \to \infty$.

Now let us prove the existence of minimum. For that, consider the set

$$K = \{X \mid F(X) \le F(0)\}$$

The function F is convex (and hence continuous), so the set K is closed. As proved above, $\lim_{X\to\infty} F(X) = +\infty$, therefore the set K is bounded.

The space $\mathcal{E}$ is finite-dimensional, the set K is compact, therefore the function F always attains minimum on K, which is its minimum on the whole space $\mathcal{E}$. □

So far, we have obtained a sufficient condition for the existence of the minimum of a convex function.

2.2 Multipartite setting

Now return to our initial setting. We are dealing with integrals over the set $\mathbf{P}$ of pure product projectors in a tensor product space $\mathbf{H} = H^{(1)} \otimes \cdots \otimes H^{(N)}$. We shall need the following:

Proposition 4 (**Multipolarization identity**). *Any quadratic form Y in $\mathbf{H}$ is completely defined by its values on $\mathbf{P}$.*

$$\langle e^{(1)} \otimes \cdots \otimes e^{(N)} | Y | f^{(1)} \otimes \cdots \otimes f^{(N)} \rangle =$$

$$= 4^{-N} \sum_{k^{(1)},\ldots k^{(N)}=0}^{3} \mathrm{i}^{(k^{(1)}+\cdots+k^{(N)})} \left\langle \bigotimes_{I=0}^{N} \left(e^{(I)} + \mathrm{i}^{k^{(I)}} f^{(I)} \right) \middle| Y \middle| \bigotimes_{I=0}^{N} \left(e^{(I)} + \mathrm{i}^{k^{(I)}} f^{(I)} \right) \right\rangle \tag{10}$$

Proof. Verified by direct calculation. □

Remark. In the sequel we shall use the following consequence of this proposition:

$$\forall \mathrm{P} \in \mathbf{P} \quad \mathrm{Tr}(X\mathrm{P}) = 0 \quad \Rightarrow \quad X = 0 \tag{11}$$

Now return to the convex function (6)

$$G(X) \; = \; \int_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \mathrm{d}\sigma(\mathrm{P}) \; - \; \mathrm{Tr}(X\rho)$$

Proposition 5. *When the function $G(X)$ has minimum, this minimum is strict.*

Proof. Suppose there are two minimal points $B_0 \neq B_1$ of $G(X)$. Denote $V = B_1 - B_0$ and consider a family $B_t = B_0 + Vt$. Consider the function

$$g(t) \;=\; G(B_t) \;=\; \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(B_t\mathrm{P})} \mathrm{d}\sigma(\mathrm{P}) \;-\; \mathrm{Tr}(B_t\rho)$$

for $t \in [0,1]$. This function is constant because G is *convex*. Therefore its second derivative vanishes, but:

$$g''(t) \;=\; \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(B_t\mathrm{P})}\,(\mathrm{Tr}(V\mathrm{P}))^2 \mathrm{d}\sigma(\mathrm{P}) \tag{12}$$

In the meantime $V \neq 0$, therefore $\int_{\mathbf{P}} (\mathrm{Tr}(V\mathrm{P}))^2 \mathrm{d}\sigma(\mathrm{P}) > 0$ due to (11), so $g''(t) > 0$ — contradiction. □

However, the *existence* of the minimum of $G(X)$ still can not be directly verified. In the meantime, the approximations of $G(X)$ by the family of convex functions

$$G_k(X) \;=\; \int_{\mathbf{P}} \left(1 + \frac{\mathrm{Tr}(X\mathrm{P})}{2k}\right)^{2k} \mathrm{d}\sigma(\mathrm{P}) \;-\; \mathrm{Tr}(X\rho) \tag{13}$$

possess the following property:

Proposition 6. *Whatever be ρ, for any k there exists minimum of the function $G_k(X)$ and this minimum is strict.*

Proof. First prove that the minimum exists. Fix an arbitrary $X \neq 0$ and apply the sufficient condition (9) by showing that $G_k(tX) \to +\infty$. Let

$$Y = \mathbb{I} + \frac{tX}{2k}$$

then

$$G_k(tX) = \int_{\mathbf{P}} \mathrm{Tr}(YP)^{2k} \mathrm{d}\sigma(\mathrm{P}) - t\ \mathrm{Tr}(XA)$$

For all $Y \neq 0$, introduce

$$N(Y) = \frac{\int_{\mathbf{P}} \mathrm{Tr}(YP)^{2k} \mathrm{d}\sigma(\mathrm{P})}{\mathrm{Tr}(Y^2)^k} \geq 0$$

Being homogeneous, $N(Y)$ is completely defined by its values on the compact set defined by $\mathrm{Tr}(Y^2) = 1$. Let us prove that $N(Y)$ is strictly positive. Suppose $N(Y) = 0$ for some Y, then $\mathrm{Tr}(YP) = 0$ for all product one-dimensional projectors $P \in \mathbf{P}$. That means, for all vectors $e^{(1)}, \ldots, e^{(N)}$

$$\langle e^{(1)} \otimes \cdots \otimes e^{(N)} | Y | e^{(1)} \otimes \cdots \otimes e^{(N)} \rangle = 0$$

Then, by virtue of (11), $Y = 0$.

Being continuous function defined on a compact set, $N(Y)$ attains its minimal value, denote it a. Then

$$G_k(tX) \geq a \,\mathrm{Tr}(Y^2)^k - t\ \mathrm{Tr}(X\rho) \ = \ at^{2k} \left(\mathrm{Tr} \left(t^{-1}\mathbb{I} + \frac{X}{2k} \right)^2 \right)^k - t\ \mathrm{Tr}(X\rho) \ \to \ +\infty$$

since $a > 0$ as $N(Y)$ was proved to be strictly positive. So the minimum of $G_k(X)$ exists.

In order prove that the minimum is strict, we proceed in a way similar to Proposition 5 with the only difference that the function $g(t)$ has the form:

$$g(t) \ = \ G_k(B_t) \ = \ \int\limits_{\mathbf{P}} \left(1 + \frac{\mathrm{Tr}(B_t \mathrm{P})}{2k} \right)^{2k} \mathrm{d}\sigma(\mathrm{P}) \ - \ \mathrm{Tr}(B_t \rho)$$

and we check its $2k$-s derivative (rather than the second one)

$$g^{(2k)}(t) \ = \ \int\limits_{\mathbf{P}} \frac{(2k-1)!}{(2k)^{2k-1}} \left(\mathrm{Tr}(V\mathrm{P})\right)^{2k} \mathrm{d}\sigma(\mathrm{P})$$

and obtain the same contradiction. □

Corollary. Whatever (separable or entangled) be ρ, it decomposes into

$$\rho \ = \ \int_{\mathbf{P}} \mathrm{P} w(\mathrm{P}) \mathrm{d}\sigma(\mathrm{P}) \tag{14}$$

with

$$w(\mathrm{P}) = \left(1 + \frac{\mathrm{Tr}(B\mathrm{P})}{2k} \right)^{2k-1} \tag{15}$$

where B is the minimal point of the function G_k, that is, $\nabla G_k(B) = 0$.

We emphasize that the obtained decomposition (14) is in general *not barycentric* as the density (15) may take negative values. Furthermore, for entangled ρ the density (15) will always take both positive and negative values. Let us consider it in more detail.

2.3 The convergence of the approximations

We start from the main equation (5)

$$\int_{\mathbf{P}} \mathrm{P}\mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \mathrm{d}\sigma(\mathrm{P}) \; - \; \rho \; = \; 0$$

which may not have solutions and replace it by a sequence of its binomial approximations

$$\int_{\mathbf{P}} \mathrm{P} \left(1 + \frac{\mathrm{Tr}(X\mathrm{P})}{2k}\right)^{2k-1} \mathrm{d}\sigma(\mathrm{P}) \; - \; \rho \; = \; 0 \tag{16}$$

each always having a unique solution. So, we have to study the conditions when these approximations turn to the main equation. For that, we explore the convergence $G_k \to G$ of functions associated with the equations in question on the set $\mathcal{E}$ of self-adjoint operators in $\mathbf{H}$.

Theorem 7. *If the function $G(X)$ has strict minimum on $\mathcal{E}$ attained in B, then the sequence B_k of the minimal points of $G_k(X)$ converge to B.*

Proof. First prove that the convergence $G_k \to G$ is uniform on any compact subset of $\mathcal{E}$. For any $0 \leq a \leq n$ direct calculation yields

$$\max_{|x|\leq a} \left| e^x - \left(1 + \frac{x}{n}\right)^n \right| = \max \left\{ e^a - \left(1 + \frac{a}{n}\right)^n , \; e^{-a} - \left(1 - \frac{a}{n}\right)^n \right\}$$

therefore

$$\left(1 + \frac{x}{n}\right)^n - \mathrm{e}^x \to 0 \tag{17}$$

uniformly on any finite interval in $\mathbb{R}$. For $n = 2k$

$$|G_k(X) - G(X)| \leq \int\limits_{\mathbf{P}} \left| \left(1 + \frac{\mathrm{Tr}(X\mathrm{P})}{2k}\right)^{2k} - \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \right| \mathrm{d}\sigma(\mathrm{P})$$

Since X ranges over a compact set, $|\,\mathrm{Tr}(X\mathrm{P})| \leq C$ for some C not depending on X. The integration set $\mathbf{P}$ is compact, so the convergence $G_k \to G$ is uniform.

Now let us prove that the sequence B_k tends to B. The minimum is strict, so $G(X) > G(B)$ for any $X \neq B$. The sphere $\|X - B\| = \varepsilon$ is compact, then there exists

$$a = \min_{\|X-B\|=\varepsilon} G(X) - G(B) > 0$$

Since $G_k \to G$ uniformly on compact sets, a number N_ε exists such that for any $k > N_\varepsilon$

$$G(X) - \frac{a}{3} \; < \; G_k(X) \; < \; G(X) + \frac{a}{3}$$

for any $X: \; \|X - B\| \leq \varepsilon$. Then for any such X

$$\frac{2}{3}a = a - \frac{a}{3} \leq G(X) - G(B) - \frac{a}{3} < G_k(X) - G(B)$$

So,

$$\frac{2}{3}a + G(B) < G_k(X) \qquad \forall X: \; \|X - B\| \leq \varepsilon \tag{18}$$

It remains to check that $\|B_k - B\| \leq \varepsilon$ for all $k > N_\varepsilon$. Suppose a number $k > N_\varepsilon$ exists such that the appropriate minimal point $\|B_k - B\| > \varepsilon$. For $t = \varepsilon / \|B_k - B\|$ consider the convex combination

$$E = (1-t) \cdot B + tB_k$$

then $\|E - B\| = \varepsilon$. Then it follows from (18) and the convexity of G_k that

$$\frac{2}{3}a + G(B) < G_k(E) \leq (1-t) \cdot G_k(B) + tG_k(BB_k) \leq$$

$$\leq (1-t) \cdot G_k(B) + tG_k(B) = G_k(B) < \frac{a}{3} + G(B)$$

— contradiction.

□

Theorem 8. *If the sequence B_k of the minimal points of $G_k(X)$ converge, then the function $G(X)$ has strict minimum at point $B = \lim B_k$.*

Proof. For any X

$$G_k(X) \geq G_k(B_k)$$

For any fixed X the sequence $G_k(X) \to G(X)$, so

$$G(X) = \lim G_k(X) \geq \overline{\lim}\, G_k(B_k) \tag{19}$$

Now check that

$$\lim \; G_k(B_k) = G(B)$$

The convergence $B_j \to B$ and $G_k \to G$ together with the continuity of each $G_k(X)$ imply

- $\forall k\, G_k(B_j) \to G_k(B)$
- $G_k(B) \to G(B)$

Applying Cantor diagonal method, we choose a subsequence B_{j_k} such that $G_k(B_{j_k}) \to G(B)$. For arbitrary X, Y we have

$$|G_k(X) - G_k(Y)| \leq \int\limits_{\mathbf{P}} \left| \left(1 + \frac{\mathrm{Tr}(X\mathrm{P})}{2k}\right)^{2k} - \left(1 + \frac{\mathrm{Tr}(Y\mathrm{P})}{2k}\right)^{2k} \right| \mathrm{d}\sigma(\mathrm{P}) + \|X - Y\| \cdot \|\rho\|$$

Observing that for any $x, y \in \mathbb{R}$

$$\left| \left(1 + \frac{x}{n}\right)^n - \left(1 + \frac{y}{n}\right)^n \right| \leq \mathrm{e}^{|x|+|y|} \cdot |x - y|$$

and taking into account that $\mathrm{Tr}(X\mathrm{P}) \leq \|X\|$, we get

$$|G_k(X) - G_k(Y)| \leq \left(\mathrm{e}^{\|X\|+\|Y\|} + \|\rho\|\right) \cdot \|X - Y\|$$

Using this, we have

$$|G_k(B_k) - G_k(B_{j_k})| \leq \left(\mathrm{e}^{\|B_k\|+\|B_{j_k}\|} + \|\rho\|\right) \cdot \|B_k - B_{j_k}\| \ \to \ 0$$

Therefore $G(B) = \lim G_k(B_k)$, so it follows from (19) that B is the minimal point of $G(X)$. In accordance with the Proposition 5 this minimum is strict.

□

The obtained criterion can be strengthened.

Theorem 9. *If the sequence B_k of the minimal points of $G_k(X)$ is bounded, then it converges, and the function $G(X)$ has strict minimum at $B = \lim B_k$.*

Proof. Since the space $\mathcal{E}$ is finite-dimensional, we can select its converging subsequence $B_{k_n} \to B$. Let us first show that $B = \min G(X)$. For any fixed X

$$G(X) = \lim G_{k_n}(X) \geq \overline{\lim}\, G_{k_n}(B_{k_n}) \tag{20}$$

Now check that

$$\lim\ G_{k_n}(B_{k_n}) = G(B)$$

Proceeding in a way similar to Theorem 6, we select a sub-subsequence $B_{k_{n_j}}$ such that $G_{k_n}(B_{k_{n_j}}) \to G(B)$ and get the required: the function $G(X)$ has minimum. Then, applying Theorem 5 we infer that the minimum is strict and using Theorem 7 we obtain that $B_k \to B$.

□

3 "Temperature" theorem and its consequences

The mapping

$$\mathcal{L}(X) = \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})}\, \mathrm{P}\, \mathrm{d}\sigma(\mathrm{P}) \tag{21}$$

from all self-adjoint operators to positive operators in $\mathbf{H}$ was shown (proposition 5) to be injective. However, $\mathcal{L}$ is not surjective: clearly, no pure product state can be represented this way. We shall show that the image of $\mathcal{L}$ contains *almost* all separable density operators.

3.1 Matching theorem

The decompositions of a given density operator ρ – both discrete and continuous are known to be non-unique. The following theorem shows that any continuous positive decomposition of ρ can be replaced by an exponential one, which we studied before.

Theorem 10 ("Temperature" theorem)**.** *Let ρ be a density matrix such that it can be represented in the form* (2)

$$\rho \;=\; \int_{\mathbf{P}} w(\mathrm{P}) \cdot \mathrm{P}\mathrm{d}\sigma(\mathrm{P})$$

with w being positive

$$w(\mathrm{P}) > 0$$

Then there exists such self-adjoint operator B that:

$$\rho \;=\; \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(B\mathrm{P})}\, \mathrm{P}\, \mathrm{d}\sigma(\mathrm{P})$$

Proof. Taking into account (7), it suffices to prove that the function $G(X) = \int_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})}\mathrm{d}\sigma(\mathrm{P}) - \mathrm{Tr}(X\rho)$ has minimum. Writing down ρ in the form (2), we get

$$G(X) \;=\; \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})}\mathrm{d}\sigma(\mathrm{P}) - \int\limits_{\mathbf{P}} w(\mathrm{P})\, \mathrm{Tr}(X\mathrm{P})\mathrm{d}\sigma(\mathrm{P}) \;=\; \int\limits_{\mathbf{P}} \left(\mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} - w(\mathrm{P})\, \mathrm{Tr}(X\mathrm{P})\right) \mathrm{d}\sigma(\mathrm{P}) \tag{22}$$

For every fixed $\mathrm{P} \in \mathbf{P}$ introduce a function

$$g(t) \;=\; \mathrm{e}^t - at \tag{23}$$

where

$$a = w(\mathrm{P})\,, \quad t = \mathrm{Tr}(X\mathrm{P})$$

The set $\mathbf{P}$ is compact, and the function $w(\mathrm{P})$ is continuous and positive, that is why it attains its extreme values

$$0 < m \leq w(\mathrm{P}) \leq M$$

Since $w(\mathrm{P})$ is a probability density, we have

$$0 < m \leq 1 \leq M \tag{24}$$

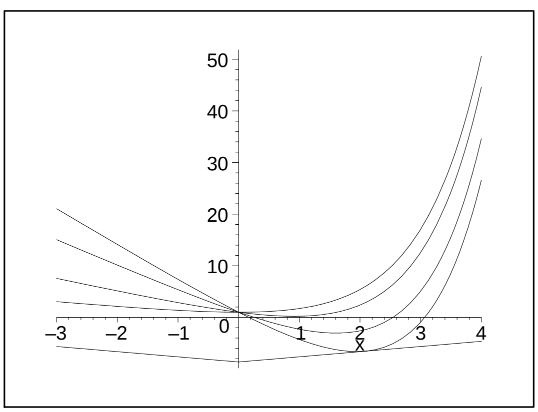
The value of a in (23) lies between m and M. The minimal value of $g(t)$ depends on the value of the parameter a as follows:

$$g_{\min} = a - a\ln a$$

By elementary (for $t \leq 0$) and routine (for $t \geq 0$) calculations we obtain

$$g(t) \geq (m+M)[1 - \ln(m+M)] + m|t| \tag{25}$$

for any real t as illustrated at this graph:



Since $t = \mathrm{Tr}(X\mathrm{P})$ and $a = w(\mathrm{P})$, this means

$$\mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} - w(\mathrm{P})\,\mathrm{Tr}(X\mathrm{P}) \;\geq\; (m+M)[1-\ln(m+M)] + m|\,\mathrm{Tr}(X\mathrm{P})|$$

Integrating this inequality over $\mathbf{P}$, we obtain the following evaluation:

$$G(X) \geq (m+M)[1-\ln(m+M)] + m\int_{\mathbf{P}} |\operatorname{Tr}(X\mathrm{P})| \mathrm{d}\sigma(\mathrm{P}) \tag{26}$$

Consider the second summand of this expression:

$$\nu(X) = \int_{\mathbf{P}} |\operatorname{Tr}(X\mathrm{P})| \mathrm{d}\sigma(\mathrm{P})$$

which is a seminorm. Proceeding like in Proposition 6 we see that $\nu(X)$ is non-degenerate. Since all norms in finite-dimensional space are equivalent, $\nu(X)$ evaluates as

$$\nu(X) \geq \mathrm{const} \cdot \|X\|$$

So far, we obtain

$$G(X) \geq (m+M)[1-\ln(m+M)] + \mathrm{const} \cdot \|X\| \to +\infty$$

Since $G(X)$ is convex, it has minimum, which is attained at certain point $B \in \mathcal{E}$. $\square$

Remark. It can be verified that the distribution $w(\mathrm{P}) = \mathrm{e}^{\operatorname{Tr}(B\mathrm{P})}$ yields the maximum for

$$S = -\int w \ln w \mathrm{d}\sigma(\mathrm{P})$$

Theorem 11 (Matching theorem)**.** *Any ρ belonging to the interior $\mathcal{D}$ of the set of separable states can be represented in the form* (4)*:*

$$\rho \ = \int\limits_{\mathbf{P}} \mathrm{e}^{\operatorname{Tr}(B\mathrm{P})} \ \mathrm{P} \, \mathrm{d}\sigma(\mathrm{P})$$

Proof. Denote by $\mathcal{M}$ the set of density operators representable as Husimi exponentials. Let $\rho_1, \rho_2 \in \mathcal{M}$ and form their convex combination $\rho = (1-t)\rho_1 + t\rho_2$. This ρ can be represented by continuous positive density $w(\mathrm{P}) = (1-t)\mathrm{e}^{\operatorname{Tr}(B_1\mathrm{P})} + t\mathrm{e}^{\operatorname{Tr}(B_2\mathrm{P})}$, therefore it follows from the "Temperature" theorem 10 that such B exists that $\rho = \int \mathrm{e}^{\operatorname{Tr}(B\mathrm{P})}\mathrm{P}\mathrm{d}\sigma(\mathrm{P})$. That means, $\mathcal{M}$ is convex together with its closure.

Let $\mathrm{P}_0 \in \mathbf{P}$ be a product pure state. Any atomic measure on a compact set $\mathbf{P}$ can be approximated by a sequence of strictly positive densities. Take such a sequence w_n on $\mathbf{P}$ such that for any continuous function f on $\mathbf{P}$

$$\lim_{n\to\infty} \int\limits_{\mathbf{P}} f(\mathrm{P}) w_n(\mathrm{P}) \ \mathrm{d}\sigma(\mathrm{P}) \ = \ f(\mathrm{P}_0)$$

Consider the sequence of operators

$$\rho_n = \int\limits_{\mathbf{P}} w_n(\mathrm{P})\ \mathrm{P}\,\mathrm{d}\sigma(\mathrm{P})$$

each belonging to $\mathcal{M}$. For any affine function $h(X)$

$$h(\rho_n) = \int\limits_{\mathbf{P}} w_n(\mathrm{P})\ h(\mathrm{P})\,\mathrm{d}\sigma(\mathrm{P}) \ \to h(\mathrm{P}_0)$$

therefore $\rho_n \to \mathrm{P}_0$. So far, the closure of $\mathcal{M}$ contains $\mathbf{P}$. The closure is shown to be convex, therefore it contains all separable density operators.

Now consider the "partition function"

$$Z(X) \ = \ \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \mathrm{d}\sigma(\mathrm{P}) \tag{27}$$

Note that for any $V \neq 0$

$$\mathrm{d}_X^2 Z(V) \ = \ \left.\frac{\mathrm{d}^2 Z(X+Vt)}{\mathrm{d}t^2}\right|_{t=0} \ = \ \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \big(\mathrm{Tr}(V\mathrm{P})\big)^2 \mathrm{d}\sigma(\mathrm{P}) > 0$$

whose positivity was established in (12). However $\mathrm{d}^2 Z = \mathrm{d}\mathcal{L}$, where $\mathcal{L} = \nabla Z$, see (21). Then the mapping $\mathcal{L}(X)$ is non-degenerate at every point $X \in \mathcal{E}$. According to the inverse image theorem, we see that $\mathcal{L}$ is local diffeomorphism, that is, each point X has a neighborhood whose image under the mapping $\mathcal{L}$ is open in $\mathcal{E}$. So, we conclude that its full image $\mathcal{L}(\mathcal{E})$ is an open subset in $\mathcal{E}$. In the meantime

$$\mathcal{M} = \mathcal{L}(\mathcal{E}) \cap \{X \in \mathcal{E} \mid \mathrm{Tr}\, X = 1\}$$

so, $\mathcal{M}$ is an open subset of the set of separable density operators: $\mathcal{M} \subseteq \mathcal{D}$.

So far, we have shown that the set $\mathcal{M}$ is an open, convex and dense subset of the set of all separable density operators, therefore $\mathcal{M}$ coincides with its interior: $\mathcal{M} = \mathcal{D}$. $\square$

3.2 Husimi coordinatization

Starting from the "partition function" Z defined above (27), introduce

$$W = \ln Z \; = \; \ln \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \mathrm{d}\sigma(\mathrm{P}) \tag{28}$$

and first calculate its gradient

$$\nabla W(X) = \frac{\nabla Z}{Z} = \frac{1}{Z} \int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}(X\mathrm{P})} \mathrm{P} \mathrm{d}\sigma(\mathrm{P})$$

Proposition 12. *The quadratic form* $\mathrm{d}^2 W$ *is nonnegatively defined and vanishes only on scalar operators.*

Proof. Calculate the value of $\mathrm{d}^2 W$ at point X on element V

$$\mathrm{d}^2 W(X;V) = \left.\frac{\mathrm{d}^2 W(X+Vt)}{\mathrm{d}t^2}\right|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left.\left(\frac{1}{Z}\int\limits_{\mathbf{P}} \mathrm{e}^{\mathrm{Tr}((X+Vt)\mathrm{P})}\, \mathrm{Tr}(V\mathrm{P}) \mathrm{d}\sigma(\mathrm{P})\right)\right|_{t=0}$$

Denoting

$$w(P) = \frac{\mathrm{e}^{\mathrm{Tr}(X\mathrm{P})}}{Z} \tag{29}$$

it reads

$$\mathrm{d}^2 W(X;V) = \int\limits_{\mathbf{P}} (\mathrm{Tr}(V\mathrm{P}))^2 w(P) \mathrm{d}\sigma(\mathrm{P}) - \left(\int\limits_{\mathbf{P}} \mathrm{Tr}(V\mathrm{P}) w(P) \mathrm{d}\sigma(\mathrm{P})\right)^2 \geq 0 \tag{30}$$

This expression vanishes if and only if $\mathrm{Tr}(V\mathrm{P}) = \mathrm{const}$ for every $\mathrm{P} \in \mathbf{P}$. All such V are the scalar operators: $V = \lambda\mathbb{I}$. $\square$

Note that for any scalar operator the function W has the property:

$$W(X + \lambda\mathbb{I}) = W(X) + \lambda$$

therefore its gradient

$$\chi(X) = \nabla W(X) \tag{31}$$

is invariant with respect to shifts along scalar operators:

$$\chi(X + \lambda\mathbb{I}) = \chi(X) \tag{32}$$

This shows that the mapping χ is not injective on $\mathcal{E}$, but its restriction to the set $\mathcal{N}$ of traceless operators from $\mathcal{E}$ becomes injective.

Theorem 13. *The mapping* (31) *establishes a diffeomorphism*

$$\chi : \mathcal{N} \to \mathcal{D} \tag{33}$$

between the set $\mathcal{N}$ of all traceless self-adjoint operators in $\mathbf{H}$ *and the interior $\mathcal{D}$ of the set of all separable density operators.*

Proof. The mapping χ was shown to be invariant with respect to shifts along scalar operators, so $\chi(\mathcal{N}) = \chi(\mathcal{E})$. The differential $\mathrm{d}\chi = \mathrm{d}^2 W$. It follows from Proposition 12 that $\mathrm{d}^2 W > 0$ on $\mathcal{N}$, therefore $\mathrm{d}\chi$ is non-degenerate on $\mathcal{N}$. So, χ is a diffeomorphism of $\mathcal{N}$ on the image $\chi(\mathcal{N})$. In turn, the image of χ coincides with the image of the mapping (21) — the density operators representable as (4). Finally, as shown in Matching theorem 11, this set coincides with $\mathcal{D}$ — the interior of the set of all separable density operators. $\square$

Concluding remarks

In this paper, the standard thermodynamical ideas were employed. However, it was not possible to replant directly the standard techniques. The first reason is that the analog of inverse temperature is no longer a scalar β, it becomes self-adjoint operator B. That is why the "Temperature theorem" needed to be reproved in the new setting. One more essential difference is the behavior of the "partition function" Z. Unlike classical case, in our setting the gradient of $\ln Z$ is invariant (32) with respect to shifts: $X \mapsto X + \lambda\mathbb{I}$.

The overall construction is general enough: we do not dwell on any particular decomposition of the state space $\mathbf{H}$ into tensor product — this notion was shown to be relative to measurements [2]. Moreover, the set of pure product states $\mathbf{P}$ could be any compact *full* set of unit vectors in the state space $\mathbf{H}$. 'Full' means that the value of any quadratic form in $\mathbf{H}$ is completely defined by its values on $\mathbf{P}$. This is the central point of the method, based on multipolarization identity (10), which was inspired by the ideas of Kôdi Husimi [3] to consider a special kind of functions defined on pure states, namely, those parametrized by values of appropriate quadratic form.

Acknowledgments. The authors acknowledge the discussion of the results offered by the participants of the Laboratory of Quantum Information (St.Petersburg, Russia). A support from research grant RFFI 0706-00119 is appreciated.

References

[1] Hukukane Nikaido, *Convex structures and economic theory*, Academic Press, New York and London (1968)

[2] Paolo Zanardi, Daniel Lidar, Seth Lloyd, *Quantum tensor product structures are observable-induced*, Phys. Rev. Lett. 92, 060402 (2004)

[3] Kôdi Husimi, *Some Formal Properties of the Density Matrix*, Proc. Phys. Math. Soc. Jpn. **22**: 264-314 (1940)